of hydrophilic amino acids. The replacements are mostly located at the periphery of the antigen binding-site but are potentially in contact with the antigen, thus encircling the center of the epitope recognized by these antibodies. These hydrophobic contacts may account for more than 15- and 10-fold increase in affinity and potency, respectively.

Intrinsically Disordered Proteins

1622-Pos Board B466

Protein Folding as a Transition Step from Ancient to the Modern Life Forms

A.K. Dunker, W.S. Chan, S.H.B. Karn, V.N. Uversky, D.J. Brooks, C.J. Oldfield, J. White, N. Perumal, P. Romero.
Indiana University, Indianapolis, IN, USA.

Several lines of evidence suggest that the first proteins on earth likely contained significantly fewer than the modern 20 amino acids. Subsequently many researchers examined the evolution of the genetic code using different criteria. Trifonov combined 40 different of these single-factor criteria into a consensus and proposed the following temporal order of addition for the amino acids: G/A, V/D, P, S, E/L, T, R, N, K, Q, I, C, H, F, M, Y, W. Brooks and co-workers estimated the amino acid composition for the Last Universal Ancestor (LUA). This composition was depleted and enriched in several of Trifonov's modern and ancient amino acids, respectively. The Brooks and coworker set of ancient proteins contains two almost equal-sized subsets: RNA-associated proteins and enzymes. We found the RNA-associated proteins from the LUA to be much more extensively depleted in the modern amino acids than were the enzymes. We also noticed that the more ancient amino acids are predominantly disorder-promoting while the more modern amino acids are predominantly order-promoting. Two different disordered protein predictors suggest the RNA-associated proteins to be disordered and the enzymes to be structured, which agrees with laboratory experiments on the modern protein counterparts. If the RNA-associated proteins are representative of the proteins present in the earliest life forms, then these proteins lacked regular 3D structure. Therefore, we propose that: 1. the change from ancient to modern life forms depended on the evolution a protein disorder-to-structure transition, thus enabling the formation of protein enzymes; and 2. this evolutionary disorder-to-order transition was enabled by the introduction of the structure-promoting amino acids during the modernization of the genetic code.

1623-Pos Board B467

Large-scale Analysis of Thermo-stable, Mammalian Proteins Provides Insights into the Intrinsically Disordered Proteome

Charles A. Galea[1], Anthony High[1], John C. Obenauer[1], Ashutosh Mishra[1], Cheon-Gil Park[1], Marco Punta[2], Avner Schlessinger[2], Jing Ma[1], Buckhard Rost[2], **Steve Otieno[1]**, Clive A. Slaughter[1], Richard W. Kriwacki[1,3].
[1]St Jude Children's Research Hospital, Memphis, TN, USA, [2]Columbia University, New York, NY, USA, [3]The University of Tennessee Health Science Center, Memphis, TN, USA.

Intrinsically disordered proteins (IDPs) are predicted to be highly abundant and play broad biological roles in eukaryotic cells, including signaling and regulation. However, these concepts are based on *in silico* analyses of whole genome sequences, not on large-scale proteomics analyses of living cells. Therefore, whether these concepts broadly apply to expressed proteins is currently unknown. Previous studies have shown that heat-treatment of cell extracts leads to partial enrichment of IDPs. Based on this, we sought to address the current dearth of knowledge about expressed IDPs by performing a large-scale proteomics study of thermo-stable proteins from mouse fibroblasts. Using a novel MudPIT strategy, we identified a total of 1,320 thermo-stable proteins from these cells and used bioinformatics methods to analyze their structural and biological properties. Interestingly, >900 of these expressed proteins were predicted to be IDPs. Unexpectedly, computational structural analyses revealed that, 1) disordered domains and coiled-coil domains occurred together in a large number of disordered proteins, suggesting functional interplay between these domains, and 2) >170 proteins contained lengthy domains (>300 residues) known to be folded. Reference to Gene Ontology Consortium functional annotations revealed that, while IDPs do, in fact, play diverse biological roles in mouse fibroblasts, they exhibit heightened involvement in particular functional categories, including, cytoskeletal structure and cell movement, metabolic and biosynthetic processes, organelle structure, cell division, gene transcription, and ribonucleoprotein complexes. We envision that these results not only reflect the specialized physiology of fibroblast cells, but also the general properties of the mouse intrinsically disordered proteome (IDP-ome). We will present these and our continuing studies of expressed, mouse IDPs, including, for example, analyses of the functional pathways associated with the over-represented functional categories noted above.
1

1624-Pos Board B468

A Robust Approach for Analyzing a Heterogeneous Structural Ensemble

Gary Daughdrill[1], David Lowry[2], Andrew Hausrath[3].
[1]University of South Florida, Tampa, FL, USA, [2]University of Idaho, Moscow, ID, USA, [3]University of Arizona, Tucson, AZ, USA.

Intrinsically unstructured proteins (IUPs) are widespread in eukaryotes and participate in numerous cellular processes, but a structural explanation of the mechanisms they employ to recognize and bind their diverse targets has proved elusive. Transcriptional activator domains (TADS) are one class of IUPs that function by recruiting other factors into transcription complexes. TADS utilize electrostatic interactions to recognize binding partners, but it is unclear how an unstructured protein could perform this activity. To investigate this question, principal component analysis was performed on the atomic contact maps of an experimentally restrained ensemble of human p53TAD. This analysis permitted the identification of persistent structural features and their relative probabilities. Thirteen clusters of structures were identified that represented 98% of the ensemble. Potential surfaces of the aligned clusters showed the negative charges of the highly acidic p53TAD are uniformly organized on one face of the clusters. This observation provides a structural basis for the recruitment of other factors into transcription complexes and further supports the hypothesis that IUPs have evolved under selection to maintain specific structural features.

1625-Pos Board B469

Unfoldomics of Human Genetic Diseases

U. Midic[1], C.J. Oldfield[2], A.K. Dunker[2], Z. Obradovic[1], **V.N. Uversky[2]**.
[1]Temple University, Philadelphia, PA, USA, [2]Indiana University School of Medicine, Indianapolis, IN, USA.

Intrinsically disordered proteins lack stable structure under physiological conditions, yet carry out many crucial biological functions, especially functions associated with regulation, recognition, signaling and control. Recently, human genetic diseases and related genes were organized into a bipartite graph (Goh, K. I., Cusick, M. E., Valle, D., Childs, B., Vidal, M., and Barabasi, A. L. (2007) The human disease network. Proc Natl Acad Sci U S A 104, 8685-90). This diseaseome network revealed several significant features such as the common genetic origin of many diseases. We analyzed the abundance of intrinsic disorder in these diseaseome network proteins by means of several prediction algorithms, and we analyzed the functional repertoires of these proteins based on prior studies relating disorder to function. Our analyses revealed that (i) Intrinsic disorder is common in proteins associated with many human genetic diseases; (ii) Different disease classes vary in the IDP contents of their associated proteins; (iii) Molecular recognition features, which are relatively short loosely structured protein regions within mostly disordered sequences and which gain structure upon binding to partners, are common in the diseasome, and their abundance correlates with the intrinsic disorder level; (iv) Some disease classes have a significant fraction of genes affected by alternative splicing, and the alternatively spliced regions in the corresponding proteins are predicted to be highly disordered; and (v) Correlations were found among the various diseasome graph-related properties and intrinsic disorder. These observations provide the basis for the construction of the human-genetic-disease-associated unfoldome.

1626-Pos Board B470

Modeling the Unfolded States of Tau protein and p21(145-164)

Austin V. Huang.
MIT, Somerville, MA, USA.

Intrinsically disordered proteins (IDPs) play essential roles in a number of normal and pathological processes, but unlike most other proteins they can adopt a variety of distinct conformations in solution. Here we propose a novel approach, called Energy-minima Mapping and Weighting (EMW), for constructing models of IDPs. The method samples energetically favorable conformations within an IDP and uses these structures to construct ensembles that are consistent with a given set of experimental data. A unique feature of the method is that it does not strive to generate a single ensemble that represents the unfolded state. Instead we construct a number of candidate ensembles, each of which agrees with a given set of experimental constraints (such as NMR chemical shifts and hydrodynamic radii and residual dipolar coupling constants) and focus our analysis on local structural features that are present in all of the independently generated ensembles. We apply the method to two natively unfolded proteins: tau protein, which plays a role in Alzheimer's Disease pathology, and ($p21^{145-164}$), a small IDP that binds to approximately 25 targets and is believed to play a role in cellular signaling. For tau protein, we deduce structural features that may explain the proclivity of tau mutants to form pathologic aggregates and in the case of p21, we demonstrate that the peptide's intrinsic